

# The VIMOS Public Extragalactic Redshift Survey (VIPERS)

## Unbiased clustering estimate with VIPERS slit assignment<sup>★</sup>

F. G. Mohammad<sup>1,2</sup>, D. Bianchi<sup>3</sup>, W. J. Percival<sup>4,5,3</sup>, S. de la Torre<sup>10</sup>, L. Guzzo<sup>2,1</sup>, B. R. Granett<sup>1,2</sup>, E. Branchini<sup>7,8,9</sup>, M. Bolzonella<sup>13</sup>, B. Garilli<sup>14</sup>, M. Scodeggio<sup>14</sup>, U. Abbas<sup>15</sup>, C. Adami<sup>10</sup>, J. Bel<sup>6,1</sup>, D. Bottini<sup>14</sup>, A. Cappi<sup>13,16</sup>, O. Cucciati<sup>11,13</sup>, I. Davidzon<sup>10,13</sup>, P. Franzetti<sup>14</sup>, A. Fritz<sup>14</sup>, A. Iovino<sup>1</sup>, J. Krywult<sup>17</sup>, V. Le Brun<sup>10</sup>, O. Le Fèvre<sup>10</sup>, K. Małek<sup>18,10</sup>, F. Marulli<sup>11,12,13</sup>, M. Polletta<sup>14,19,20</sup>, A. Pollo<sup>18,21</sup>, L. A. M. Tasca<sup>10</sup>, R. Tojeiro<sup>22</sup>, D. Vergani<sup>23</sup>, A. Zanichelli<sup>24</sup>, S. Arnouts<sup>10,25</sup>, J. Coupon<sup>26</sup>, G. De Lucia<sup>27</sup>, O. Ilbert<sup>10</sup>, L. Moscardini<sup>11,12,13</sup>, and T. Moutard<sup>28</sup>

(Affiliations can be found after the references)

Received 13 July 2018 / Accepted 6 August 2018

### ABSTRACT

The VIPERS galaxy survey has measured the clustering of  $0.5 < z < 1.2$  galaxies, enabling a number of measurements of galaxy properties and cosmological redshift-space distortions (RSD). Because the measurements were made using one-pass of the VIMOS instrument on the Very Large Telescope (VLT), the galaxies observed only represent approximately 47% of the parent target sample, with a distribution imprinted with the pattern of the VIMOS slitmask. Correcting for the effect on clustering has previously been achieved using an approximate approach developed using mock catalogues. Pairwise inverse probability (PIP) weighting has recently been proposed to correct for missing galaxies, and we apply it to mock VIPERS catalogues to show that it accurately corrects the clustering for the VIMOS effects, matching the clustering measured from the observed sample to that of the parent. We then apply PIP-weighting to the VIPERS data, and fit the resulting monopole and quadrupole moments of the galaxy two-point correlation function with respect to the line-of-sight, making measurements of RSD. The results are close to previous measurements, showing that the previous approximate methods used by the VIPERS team are sufficient given the errors obtained on the RSD parameter.

**Key words.** cosmology: observations – large-scale structure of Universe – galaxies: high-redshift – galaxies: statistics

## 1. Introduction

The clustering of galaxies within galaxy surveys provides a wealth of astrophysical information, allowing measurements of galaxy formation, galaxy evolution, and cosmological parameters. Missing galaxies within surveys can however distort the clustering compared to that of the full population of the type of objects to be observed if the missed galaxies are not randomly chosen but instead cluster in a different way to the full population. Such a situation is often induced by the mechanics of the experimental apparatus, which, given a parent population of targets, limits what can actually be observed. In this paper we consider missing galaxies in the VIPERS survey (Guzzo et al. 2014; Scodeggio et al. 2018). VIPERS collected 89022 galaxy redshifts over an overall area of  $23.5 \text{ deg}^2$ , covering the W1 and W4 fields of the Canada-France-Hawaii Telescope Legacy

Survey Wide (CFHTLS-Wide)<sup>1</sup>. A colour pre-selection was used to remove galaxies at  $z < 0.5$ , helping to bring the sampling efficiency to 47%. VIPERS conducted observations using the VIMOS multi-object spectrograph (Le Fèvre et al. 2003), which applies a slit-mask to select targets for follow-up spectroscopy. A brief description of VIPERS is provided in Sect. 2.

The requirement that spectra taken with VIMOS should not overlap on the focal plane limits the placement of slits, and consequently the galaxies that can be observed. This effect is stronger along the dispersion direction compared to across it, because of the rectangular nature of the projected spectra. The occulted region around each galaxy is imprinted on the statistical distribution of the observed galaxies. There are no overlapping observations, such as those present in the Baryon Oscillation Spectroscopic Survey (BOSS, Dawson et al. 2016), meaning that the lost information cannot be recovered: we simply do not have clustering information on scales smaller than the minimum separation perpendicular to the dispersion direction. On larger scales, the slitmask still impacts on the measured clustering through the large-scale pattern imprinted on the sky, and the density dependence of the selection.

Bianchi & Percival (2017) and Percival & Bianchi (2017) presented a new method to correct for missing galaxies in surveys. This builds up a probability for each pair of galaxies in the observed sample to have been observed in a set of

<sup>★</sup> Based on observations collected at the European Southern Observatory, Cerro Paranal, Chile, using the Very Large Telescope under programs 182.A-0886 and partly 070.A-9007. Also based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada-France-Hawaii Telescope (CFHT), which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. This work is based in part on data products produced at TERAPIX and the Canadian Astronomy Data Centre as part of the Canada-France-Hawaii Telescope Legacy Survey, a collaborative project of NRC and CNRS. The VIPERS web site is <http://www.vipers.inaf.it/>

<sup>1</sup> <http://www.cfht.hawaii.edu/Science/CFHTLS/>



realisations of the survey<sup>2</sup>. These realisations, drawn from the same underlying parent catalogue, are all equally likely. Each sample can be obtained by simply re-running the targeting algorithm after moving or rotating the parent sample, or changing any random selection performed by the selection algorithm. We observe one of these sets of galaxies, and by inverse weighting by the pairwise probability of observation we force the clustering of the one realisations to match that of the set as a whole. Provided there are no pairs of zero weight, this weighting leads to a clustering estimate of the observed sample that is unbiased compared to that of the full parent sample. The method is described in more detail in Sect. 4.1.

In this paper we apply this method to remove the effects of the VIMOS slitmask from the VIPERS survey. The slitmask has a strong effect, leading to an observed clustering signal that is very different from that expected (de la Torre et al. 2013). In previous VIPERS papers this was approximately corrected using a target sampling rate (TSR) given by the fraction of potential targets placed behind a slit in a rectangular region around each targeted galaxy (Pezzotta et al. 2017). A further correction that up-weights galaxy pairs by the ratio  $[1 + w_s(\theta)]/[1 + w_p(\theta)]$  of the angular clustering of the observed  $w_s$  and parent  $w_p$  samples (de la Torre et al. 2013) was also used to improve the small-scale clustering measurements. While similar in principle to the method of Bianchi & Percival (2017) and Percival & Bianchi (2017), this relies on the missed pairs being statistically identical to the population as a whole. This is not the case in VIPERS as galaxies are more likely to be missed in denser regions where they have different properties. The TSR up-weighting method was extensively tested in past VIPERS analyses to provide a sub-percentage-level accuracy on the clustering measurements in mock catalogues. However, the TSR weighting is a parametric method that was calibrated on mock catalogues to minimise the systematic bias of the clustering measurements. It does not take into account possible differences in the clustering of simulated and observed datasets. The pairwise inverse probability (PIP) weighting scheme uses the data themselves to infer the selection probabilities providing the same level of accuracy. In this sense the new correction method is self-contained and more robust than the method based on using the TSR.

To optimise the design of the slitmasks, VIPERS uses the so-called SPOC algorithm (Slit Positioning and Optimisation Code), within the ESO VIMOS mask preparation software VMPS (Bottini et al. 2005). SPOC was designed to obtain the most spectra possible given an input parent sample. Rather than trying to change the internal properties of SPOC to make our set of realisations of the survey, we instead rely on spatially moving the survey mask and rotating the sample. We still miss all pairs that have a separation that is less than the minimum slit separation scale, but this is not an issue as we only consider larger scales here.

We use mock catalogues of VIPERS to test the new algorithm in Sect. 6, showing that it works as expected. Having corrected for the slit-mask effects, we consider how this changes the redshift-space distortions (RSD) signal within the sample. VIPERS was designed with RSD as one of the key measurements to be made: RSD are caused by the peculiar velocities of galaxies, which systematically distort redshifts leaving an

enhanced clustering signal along the line-of-sight (Kaiser 1987). By measuring the clustering anisotropy around the line-of-sight through observations of the multipole moments of the correlation function one can constrain the growth rate of cosmological structure parameterised by  $f\sigma_8$ , which constitutes the first-order contribution to the RSD signal.

Early RSD measurements from VIPERS were based on the Public Data Release 1 sample (Garilli et al. 2014), measuring  $f\sigma_8(z = 0.8)$  (de la Torre et al. 2013). Subsequent measurements from the final data sample, Public Data Release 2 (PDR2, Scoddeggio et al. 2018), were presented by Pezzotta et al. (2017). Extensions to these measurements include a configuration space joint analysis of RSD and weak-lensing (de la Torre et al. 2017), and an analysis splitting the sample based on galaxy type in order to extract extra information by comparing samples that trace the dark matter field in different ways (Mohammad et al. 2018).

We present RSD measurements made by the “standard” two-point correlation function-based method in Sect. 8. These are compared to the previous VIPERS measurements, and we show that previous slit-mask-correction techniques were sufficient to make these measurements from VIPERS. This is discussed further in Sect. 9.

To analyse the VIPERS-PDR2 data we used the same fiducial cosmology adopted in previous VIPERS clustering analyses, that is, a flat  $\Lambda$ CDM cosmology with parameters  $(\Omega_b, \Omega_m, h, n_s, \sigma_8) = (0.045, 0.3, 0.7, 0.96, 0.80)$ .

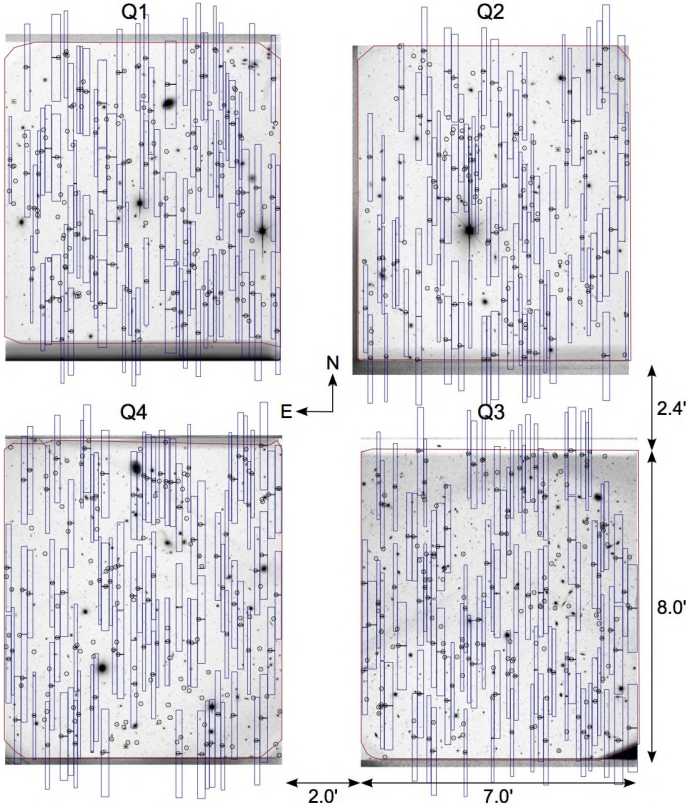
## 2. The VIPERS survey

The VIPERS survey extends over an area of  $23.5 \text{ deg}^2$  within the W1 and W4 fields of the CFHTLS-Wide. The VIMOS multi-object spectrograph (Le Fèvre et al. 2003) was used to cover these two fields with a mosaic of 288 pointings, 192 in W1 and 96 in W4. Given the VIMOS footprint, which consists of four distinct quadrants separated by an empty “cross” of about 2 arcmin width (see Fig. 1), the survey area includes a regular grid of gaps where no galaxies were observed (see following section). Target galaxies were selected from the CFHTLS-Wide catalogue to a faint limit of  $i_{AB} = 22.5$ , applying an additional  $(r - i)$  versus  $(u - g)$  colour preselection that efficiently and robustly removes galaxies at  $z < 0.5$ . Coupled with a highly optimised observing strategy (Scoddeggio et al. 2009), this doubles the mean galaxy sampling efficiency in the redshift range of interest compared to a purely magnitude-limited sample, bringing it to 47%.

Spectra were collected at moderate resolution ( $R \simeq 220$ ) using the LR Red grism, providing a wavelength coverage of 5500–9500 Å. The typical redshift error for the sample of reliable redshifts is  $\sigma_z = 0.00054(1 + z)$ , which corresponds to an error on a galaxy peculiar velocity at any redshift of  $163 \text{ km s}^{-1}$ . These and other details are given in the PDR-2 release paper (Scoddeggio et al. 2018). A discussion of the data reduction and management infrastructure was presented in Garilli et al. (2014), while a complete description of the survey design and target selection was given in Guzzo et al. (2014). The dataset used here is the same early version of the PDR-2 catalogue used in Pezzotta et al. (2017) and de la Torre et al. (2017), from which it differs by a few hundred redshifts revised during the very last period before the release. In total it includes 89 022 objects with measured redshifts. As in all statistical analyses of the VIPERS data, only measurements with quality flags 2–9 (inclusive) are used, corresponding to a sample with a redshift confirmation rate of 96.1% (for a description of the quality flag scheme, see Scoddeggio et al. 2018).

<sup>2</sup> With the term “survey realisation” we indicate a possible outcome of the spectroscopic observation given an underlying parent sample. It is not to be confused with the term “survey mocks” that are built from an ensemble of parent catalogues keeping the observational setup fixed.





**Fig. 1.** Example of the slit/spectrum distribution over a full VIMOS pointing, showing the disposition of the four quadrants and the “cross” among them. The circles identify the targets selected by the SPOC optimisation algorithm. The elongated blue rectangles reproduce the “shadow” of the 2D spectrum that will result from each target in the final spectroscopic exposure. The thin red lines show the boundary of the actual spectroscopic mask, traced pointing-by-pointing through an automatic detection algorithm that follows the borders of the illuminated area (see Guzzo et al. 2014, for details)

The procedures for defining the target list within the VIMOS spectroscopic masks were described in detail in Bottini et al. (2005). Within the VMPS environment, the SPOC algorithm is used to optimise the position, size and, total number of slits. The final solution is derived by cross-correlating the user target catalogue with the corresponding object positions in a VIMOS direct exposure of the field (“pre-image”), observed beforehand. This operation matches the astrometric coordinates to the actual instrument coordinate system, selecting the subset of objects that will eventually deliver a spectrum and, potentially, a redshift measurement.

SPOC aims at finding an optimal disposition of the slits, packing the largest possible number of spectra over each quadrant (see Bottini et al. 2005 for a detailed description of SPOC). This happens irrespectively of the parent sample angular clustering. As such, it will tend to build a distribution that is more homogeneous on the sky compared to the full galaxy population at the corresponding magnitude limit. The denser the parent galaxy sample, the stronger the bias. If the number density of galaxies on the sky is much larger than the maximum density of slits that can be packed, SPOC will essentially pick galaxies in a regular grid, packing the spectra in regular rows on top of each other. This is not quite the case for VIPERS, for which the relatively bright magnitude limit allows for targeting, on average, about one half of the available galaxies, as shown in Fig. 1. In this way, the measured sample still preserves a significant fraction of

the original angular clustering. Still, a bias is inevitably introduced and needs to be properly accounted for in any clustering measurement, which is the subject of this paper. In addition, the finite size of slits introduces a proximity effect that also needs to be corrected for when computing galaxy clustering.

Figure 1 shows an example VIMOS observation. The overall mosaic of such pointings composing the full VIPERS survey is shown in Fig. 2 for the two survey areas, W1 and W4. The boundaries of each single observation are described by the black polygons. In this figure, galaxies in the photometric parent sample and in the final VIPERS-PDR2 redshift catalogue are over-plotted as red and blue dots, respectively. The gaps of the VIMOS footprint are clearly visible as vertical and horizontal stripes, in which only unobserved objects, marked in red, are present. In addition, the overall survey mask includes: (a) gaps in the photometric sample due to bright star or photometric problems (small irregular empty regions); (b) fully failed quadrants due to mechanical failure in the VIMOS metal mask insertion before the spectroscopic observation (white regular rectangles, mostly in W4); and (c) specific details in the spectroscopic observations, such as, for example, vignetting by the VLT guide probe (described by the red line in Fig. 1; see Guzzo et al. 2014; Scodreggio et al. 2018, for details).

Throughout this work we have defined, as parent catalogue, the photometric catalogue selected according to the VIPERS target selection function (Guzzo et al. 2014), including all galaxies matching the external boundaries of the VIPERS-PDR2 sample, but with no mask applied. We have also ascribed the empty pointings and quadrants in the VIPERS-PDR2 sample to the photometric mask to avoid unnecessary complications in the implementation of the pipeline used for this analysis.

### 3. VIPERS Mocks

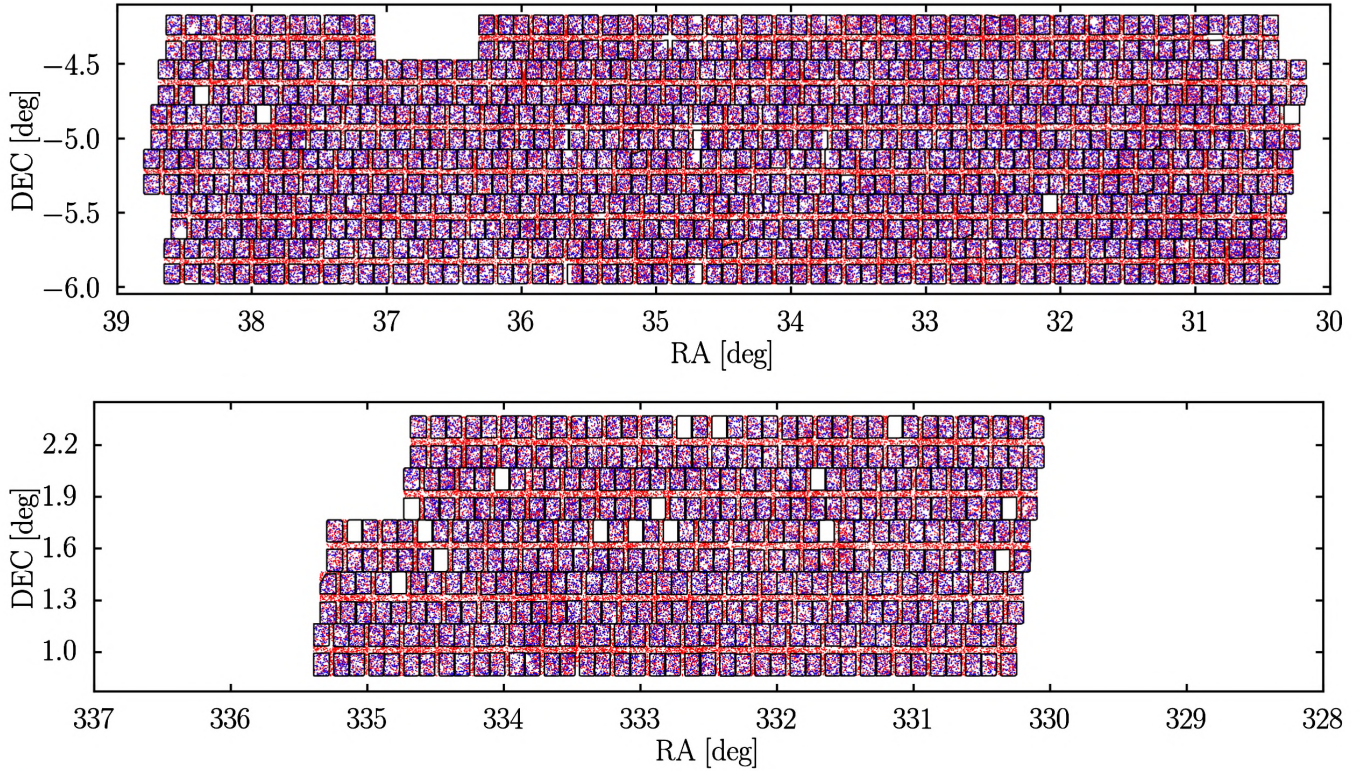
VIPERS mocks are based on the Big MultiDark Planck (BigMDPL, Klypin et al. 2016) dark matter N-body simulation. The simulation was carried out in the flat  $\Lambda$ CDM cosmological model with parameters:  $(\Omega_m, \Omega_b, h, n_s, \sigma_8) = (0.307, 0.048, 0.678, 0.96, 0.823)$ . Since the resolution is not sufficient to match the typical halo masses probed by VIPERS, low-mass haloes were added following the recipe proposed by de la Torre & Peacock (2013).

Dark-matter halos were populated with galaxies using halo occupation distribution prescriptions with parameters calibrated using luminosity-dependent clustering measurements from early VIPERS data. We refer the reader to de la Torre et al. (2013, 2017) for a detailed description of the procedure.

We used a set of 153 independent realistic parent and VIPERS-like mocks for each of the two VIPERS fields, W1 and W4. VIPERS-like mocks were obtained from the corresponding set of parent mocks in two steps: first, VIPERS targeting algorithm was applied by means of SPOC using the grid of VIPERS pointings; afterwards the footprint of VIPERS spectroscopic and photometric masks was imprinted to include the effect of obscured sky regions and quadrant vignetting (see Sect. 2). We also included the effect of VIPERS redshift error in the mock catalogues by blurring the cosmological redshifts using a Gaussian distribution of width  $\sigma_z/(1+z) = 0.00047$ . Although different from the latest estimate from the PDR2 data, we used this value to perform a fair comparison of our results with those in Pezzotta et al. (2017).

We used this set of mock samples to test the reliability of the weighting schemes proposed in Bianchi & Percival (2017) and Percival & Bianchi (2017). The same set of mocks was also





**Fig. 2.** Scatter plot in the (RA, Dec) plane for galaxies in the parent sample (red dots) and VIPERS-PDR2 catalogue (blue dots). *Top and bottom panels:* W1 and W4 fields, respectively. Portions of the sky unobserved in the spectroscopic samples due to defects in the photometric sample, bright stars, or missing quadrants have been ascribed to the photometric mask.

employed to estimate the data covariance matrix and quantify the systematic bias on estimates of the growth rate of structure.

#### 4. Measurements

We measured the anisotropic two-point correlation function  $\xi(s, \mu)$  as a function of the angle-averaged pair separation  $s$  and the cosine  $\mu$  of the angle between the pair separation and the line of sight. We employed the minimum variance Landy–Szalay estimator (Landy & Szalay 1993),

$$\xi(s, \mu) = \frac{DD(s, \mu) - 2 DR(s, \mu)}{RR(s, \mu)} + 1, \quad (1)$$

where DD, DR, and RR are the data-data, data-random, and random-random normalized pair counts, respectively. We binned  $\mu$  in 200 linear bins in the range  $0 \leq \mu \leq 1$  taking the mid-point of each bin as reference. The pair separation  $s$  is instead binned using logarithmic bins,

$$\log s_{i+1} = \log s_i + \Delta s_{\log}, \quad (2)$$

with  $\Delta s_{\log} = 0.1$ . The measurement in each pair separation bin is referenced to the logarithmic mean,

$$\log \langle s_i \rangle = \frac{\log s_i + \log s_{i+1}}{2}. \quad (3)$$

The multipole moments  $\xi^{s,(\ell)}(s)$  of the two-point correlation function are defined as its projection on the Legendre polynomials  $L_\ell(\mu)$ . Since we deal with discrete bins of the variable  $\mu$ , we replaced the integral by the Riemann sum such that,

$$\xi^{s,(\ell)}(s_i) = (2\ell + 1) \sum_{j=1}^{200} \xi^s(s_i, \mu_j) L_\ell(\mu_j) \Delta\mu. \quad (4)$$

When performing the angular pair counts  $DD_a(\theta)$  and  $DR_a(\theta)$  we used 100 linear bins within  $0^\circ \leq \theta \leq 8^\circ$ . This range is sufficiently large to cover a transverse pair separation of  $\sim 185 h^{-1} \text{ Mpc}$  at  $z = 0.5$  in VIPERS fiducial cosmology.

Following Pezzotta et al. (2017) we divided the redshift range  $0.5 < z < 1.2$  covered by VIPERS into two bins spanning  $0.5 < z < 0.7$  and  $0.7 < z < 1.2$  with effective redshifts of  $z_{\text{eff}} = 0.60$  and  $z_{\text{eff}} = 0.86$ , respectively. The subsample at low redshifts contains 30 910 galaxies while the one at high redshifts includes 33 679 galaxies. These parameters are listed in Table 1. Since VIPERS targeting over W1 and W4 fields was performed using the same observational setup we treated them as a single survey and performed the pair counts simultaneously on both fields rather than combining the measurements of the correlation function from each field.

##### 4.1. Mitigating for missing targets

The PIP approach provides us with unbiased estimates of the galaxy pair counts in the presence of missing observations, with the only formal requirement being that no pair has zero probability of being observed (Bianchi & Percival 2017).

At each separation  $s$ , the data-data pair counts are obtained as

$$DD(s) = \sum_{\mathbf{x}_m - \mathbf{x}_n \approx s} w_{mn} \frac{DD_a^{(p)}(\theta)}{DD_a(\theta)}, \quad (5)$$

where  $w_{mn} = 1/p_{mn}$  is the inverse of the selection probability of the pair formed by the galaxies  $m$  and  $n$ , whereas  $DD_a^{(p)}$  and  $DD_a$  represent the angular pair counts of parent and observed sample, respectively. The observed angular pair counts are, in



**Table 1.** Parameters characterising the two VIPERS subsamples split by redshift as used in this work.

Redshift	$N_{\text{gal}}$	$V$ [ $\text{h}^{-3}\text{Gpc}^3$ ]	$z_{\text{eff}}$
$0.5 < z < 0.7$	30,910	$1.76 \times 10^{-3}$	0.60
$0.7 < z < 1.2$	33,679	$7.34 \times 10^{-3}$	0.86

**Notes.** “Redshift” denotes the redshift range,  $N_{\text{gal}}$  is the number of galaxies,  $V$  stands for the volume in the VIPERS fiducial cosmology (see Sect. 1) and  $z_{\text{eff}}$  is the effective redshift of the sample computed as the median of the mean redshifts of galaxy-galaxy pairs with separations  $5 \text{ h}^{-1} \text{ Mpc} < s < 50 \text{ h}^{-1} \text{ Mpc}$ . All figures refer to the full VIPERS, that is, both W1 and W4 fields.

turn, computed via the same  $w_{mn}$  weights,

$$DD_a(\theta) = \sum_{\mathbf{u}_m, \mathbf{u}_n \approx \cos(\theta)} w_{mn}. \quad (6)$$

For brevity, we have adopted the notation  $\sum_{\mathbf{x}_m - \mathbf{x}_n \approx s}$  and  $\sum_{\mathbf{u}_m, \mathbf{u}_n \approx \cos(\theta)}$ , with  $\mathbf{u}_i = \mathbf{x}_i/|\mathbf{x}_i|$ , to indicate that the sum is performed in bins of  $s$  and  $\theta$ , respectively. Similarly, for the data-random pair counts,

$$DR(s) = \sum_{\mathbf{x}_m - \mathbf{y}_n \approx s} w_m \frac{DR_a^{(p)}(\theta)}{DR_a(\theta)}, \quad (7)$$

where  $w_m = 1/p_m$  is the inverse of the selection probability of the galaxy  $m$ , and

$$DR_a(\theta) = \sum_{\mathbf{u}_m, \mathbf{u}_n \approx \cos(\theta)} w_m. \quad (8)$$

We evaluate the selection probabilities  $p_{mn}$  and  $p_m$  empirically, by creating an ensemble of possible outcomes of the target selection given an underlying parent catalogue; that is, we rerun the slit-assignment algorithm on the same parent sample several times (see Sect. 5). As discussed in Bianchi & Percival (2017), rather than storing all the PIP weights (one for each pair), it is convenient to compress the information in the form of individual bitwise weights (one for each galaxy). The bitwise weight of a galaxy  $w_i^{(b)}$  is defined as a binary array, of length  $N_{\text{runs}}$ , in which the  $n$ -th bit equals 1 if the galaxy has been selected in the  $n$ -th targeting realisation and 0 otherwise.  $N_{\text{runs}}$  represents, by construction, the total number of realisations. For convenience, we use base-ten integers to encode the bitwise weights. The PIP weights are obtained “on the fly”, while doing pair counts, as

$$w_{mn} = \frac{N_{\text{runs}}}{\text{popcnt}[w_m^{(b)} \& w_n^{(b)}]}, \quad (9)$$

where  $\&$  and  $\text{popcnt}$  are fast bitwise operators, which multiply two integers bit by bit and return the sum of the bits of the resulting integer, respectively. Similarly, for individual weights, we have

$$w_m = \frac{N_{\text{runs}}}{\text{popcnt}[w_m^{(b)}]}. \quad (10)$$

The requirement that all pairs are observable (they can be observed in at least one VIPERS realisation) means that the expectation value of the PIP estimator (excluding angular up-weighting) matches the clustering of all of the pairs within the

parent sample - those targeted for possible VIPERS observation. Pairs in the parent sample that cannot be observed would formally have infinite weight but, practically, they would never appear in the pair counts in a particular realisation of VIPERS<sup>3</sup>. If we have some pairs that are not observable (they have zero probability of observation), angular up-weighting can serve two different purposes:

- (i) The number of unobservable pairs is not negligible, but the clustering of these pairs is statistically equivalent to that of the observable ones. This happens when being observable or not is a property that does not depend on clustering; for example, when galaxies fall in a blind spot of the instrument’s focal plane (see Bianchi et al. 2018 for a more detailed discussion). In this case it is formally correct to use the full set of observable plus unobservable pairs to perform angular upweighting to recover unbiased estimates of the three-dimensional clustering. We note that here these regions would not be excluded in the mask used to create the random catalogue.
- (ii) The unobservable pairs are such because of their clustering but the total number is small enough that their effect is negligible, at least on the scales of interest. In this second scenario angular upweighting is simply a way to reduce the variance and the more self-consistent approach is to use only the set of observable pairs. Using the full set of pairs could potentially increase the effect of the unobservable pairs.

As discussed in Sect. 6, the VIPERS survey is compatible with category (ii). Interestingly, we find that the mean fraction of unobservable pairs in mock samples is about a factor of two larger than what is shown in Fig. 3 for the VIPERS-PDR2 galaxy sample. This points to some difference between mocks and data in terms of galaxy clustering. Unlike the weighting schemes calibrated on simulated datasets (e.g. TSR weights), PIP weights are built to be insensitive to this difference. We use mocks just to verify that the effect of unobservable pairs is confined to the smallest scales. The above mentioned factor two guaranties that the same conclusion holds for real data.

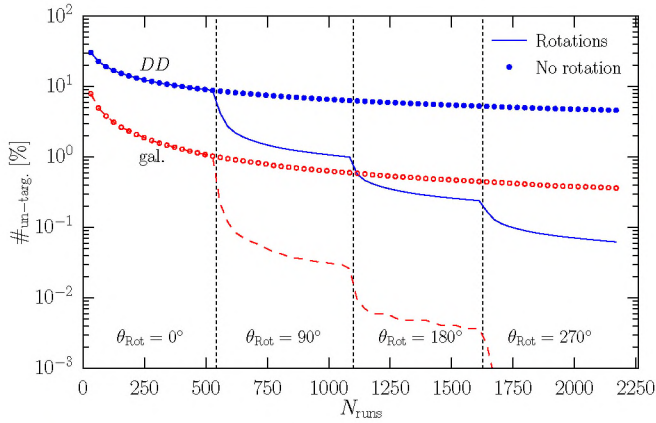
#### 4.2. Correcting for redshift failures

The reliability of each VIPERS redshift measurement is quantified by a quality flag. Spectroscopic redshift measurements with a quality flag 2–9 (inclusive) have a redshift confirmation rate of 96.1% and are regraded as reliable. We label all objects that do not satisfy this condition as “redshift failures”. The reliability of a redshift measurement depends on a number of factors such as the field-to-field observational conditions and the presence of clear spectral features and presents a correlation with some galaxy properties such as colour and luminosity. The effect of redshift failures is quantified by means of the spectroscopic success rate (SSR) defined as the ratio between the number of objects with a reliable redshift measurement (in our case the ones with a quality flag between 2 and 9) and the total number of targets placed behind a slit in a given VIMOS quadrant. It is computed as a function of the galaxy rest-frame  $U-V$  colour and  $B$ -band luminosity and is assigned to each galaxy with a reliable redshift measurement.

To correct the clustering measurements against redshift failures, we have up-weighted each galaxy by the corresponding

<sup>3</sup> For the sake of clarity, we note that  $S_{\text{pairs}} \subseteq S_{\text{observable pairs}} \subseteq S_{\text{observed pairs}}$ , where  $S_x$  stands for set of  $x$ . We also note that, in general, it is not possible to infer  $S_{\text{observable pairs}}$  from  $S_{\text{observed galaxies}}$ .





**Fig. 3.** Fraction of unobservable galaxies and pairs of galaxies in the VIPERS parent catalogue as a function of the number of targeting runs  $N_{\text{runs}}$ . Points show the case when multiple survey realisations are generated only spatially moving the spectroscopic mask, while lines result from also rotating the parent catalogue by  $\theta_{\text{rot}}$ . For the latter case, vertical dashed lines delimit the subset of targeting runs sharing the same  $\theta_{\text{rot}}$ . Blue filled points and continuous line show the fraction of unobservable galaxy-galaxy pairs while red empty markers and dashed line correspond to individual galaxies.

weight  $w_{\text{SSR}} = \text{SSR}^{-1}$ . Equations (5) and (6) are therefore modified as

$$DD(s) = \sum_{\mathbf{x}_m - \mathbf{x}_n \approx s} w_{mn} \frac{DD_a^{(p)}(\theta)}{DD_a(\theta)} \times w_{\text{SSR}}^{(m)} w_{\text{SSR}}^{(n)}, \quad (11)$$

and

$$DD_a(\theta) = \sum_{\mathbf{u}_m, \mathbf{u}_n \approx \cos(\theta)} w_{mn} \times w_{\text{SSR}}^{(m)} w_{\text{SSR}}^{(n)}. \quad (12)$$

Data-random cross-pair counts in Eqs. (7) and (8) now become,

$$DR(s) = \sum_{\mathbf{x}_m - \mathbf{y}_n \approx s} w_m \frac{DR_a^{(p)}(\theta)}{DR_a(\theta)} \times w_{\text{SSR}}^{(m)}, \quad (13)$$

and

$$DR_a(\theta) = \sum_{\mathbf{u}_m, \mathbf{u}_n \approx \cos(\theta)} w_m \times w_{\text{SSR}}^{(m)}, \quad (14)$$

respectively.

The effect of redshift failures is not reproduced in the mock catalogues. We therefore make use of spectroscopic success rates only when dealing with the VIPERS-PDR2 galaxy catalogue.

## 5. Pipeline

The weighting scheme presented in Sect. 4.1 relies on generating multiple survey realisations to assign selection probabilities and correct the pair counts. In principle, for a slit or fiber assignment scheme that randomly selects targets in the presence of collided objects, this can be achieved by simply re-running the targeting algorithm  $N_{\text{runs}}$  times on the parent catalogue, with different random selection choices each time.

As described in Sect. 2, SPOC applies a deterministic algorithm to maximise the number of slits assigned to potential targets, with no free parameters. Re-running the targeting algorithm with the same configuration of parent sample and spectroscopic mask would produce exactly the same outcome. We therefore

generated multiple realisations of the spectroscopic observations from a given parent catalogue by spatially moving the spectroscopic mask in the (RA, Dec) plane. As the VIPERS fields are equatorial, we can accurately quantify small shifts in the survey position using  $\Delta\text{RA}$  and  $\Delta\text{Dec}$ . Given the periodicity in the pattern of pointings in the VIPERS spectroscopic mask, the amount of this shift, with respect to the original VIPERS configuration, was taken as being smaller than the size of a single VIMOS pointing. We generated  $N_{\text{runs}} = 2170$  VIPERS target realisations on each parent sample. The first of these 2170 such runs was kept fixed to the actual VIPERS-PDR2 position.

The VIPERS spectroscopic mask is defined only over the area covered by the actual VIPERS observations. A shift would therefore inevitably yield galaxies at the edges of the sample to be covered by a lower number of targeting runs with respect to those located near the centre (Fig. 5). Rather than having to keep track of this, we replicated the grid of VIPERS pointings beyond the survey area such that in each run, all portions of the parent catalogue are covered by a VIMOS pointing. However, unlike the pointings in the original spectroscopic mask, we do not know the exact shapes of the quadrants belonging to the “artificial” pointings outside the survey area, so we used the shapes of the quadrants in the original VIPERS spectroscopic mask as templates and randomly assigned them to the artificial pointings. We henceforth refer to the new mask as the “extended spectroscopic mask”.

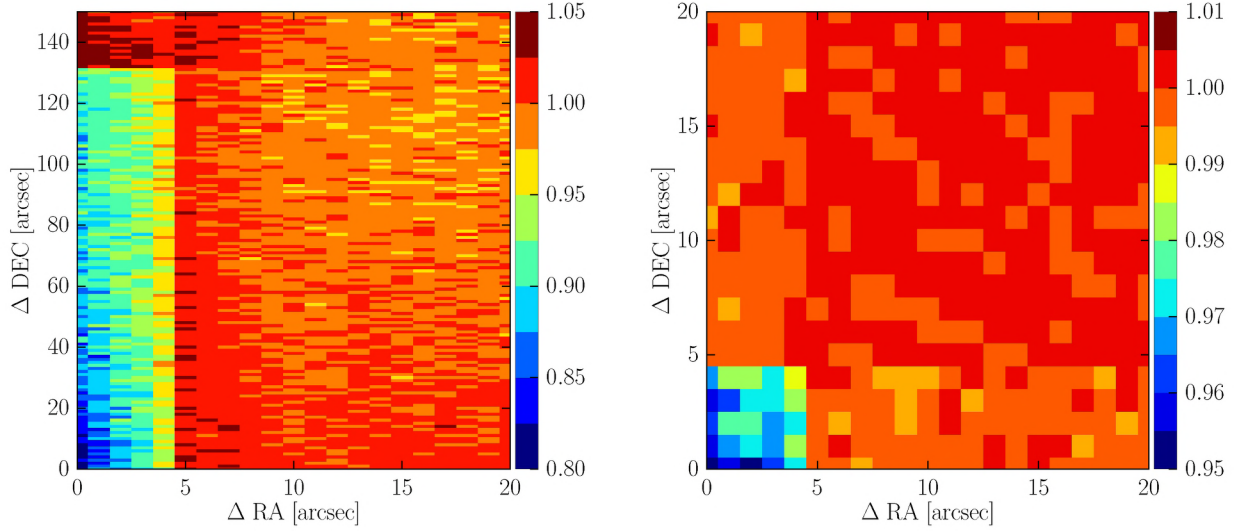
Only shifting the extended spectroscopic mask by small offsets with respect to the parent sample would require a very large number of targeting runs to accurately infer the selection probabilities and reach sub-percent level accuracy on the measurements of the two-point correlation function (see Fig. 3). In particular, after  $N_{\text{runs}} = 2170$  targeting runs obtained by only shifting the extended spectroscopic mask,  $\sim 0.6\%$  of parent galaxies remain unobserved in any of these realisations and therefore cannot be assigned a targeting probability (red empty circles in Fig. 3). This fraction increases to  $\sim 5.5\%$  for galaxy-galaxy pairs (blue filled points in Fig. 3). This is due to the fact that under particular conditions such as in very close pairs, SPOC systematically selects the same objects in different targeting runs. This effect is quantified by the 2D angular completeness function of the sub-sample of observable pairs (i.e. the ones that are observed in any of the 2170 targeting runs) in the (RA, Dec) plane,

$$C(\text{RA}, \text{Dec}) = \frac{1 + w_{\text{targ}}(\text{RA}, \text{Dec})}{1 + w_{\text{par}}(\text{RA}, \text{Dec})}, \quad (15)$$

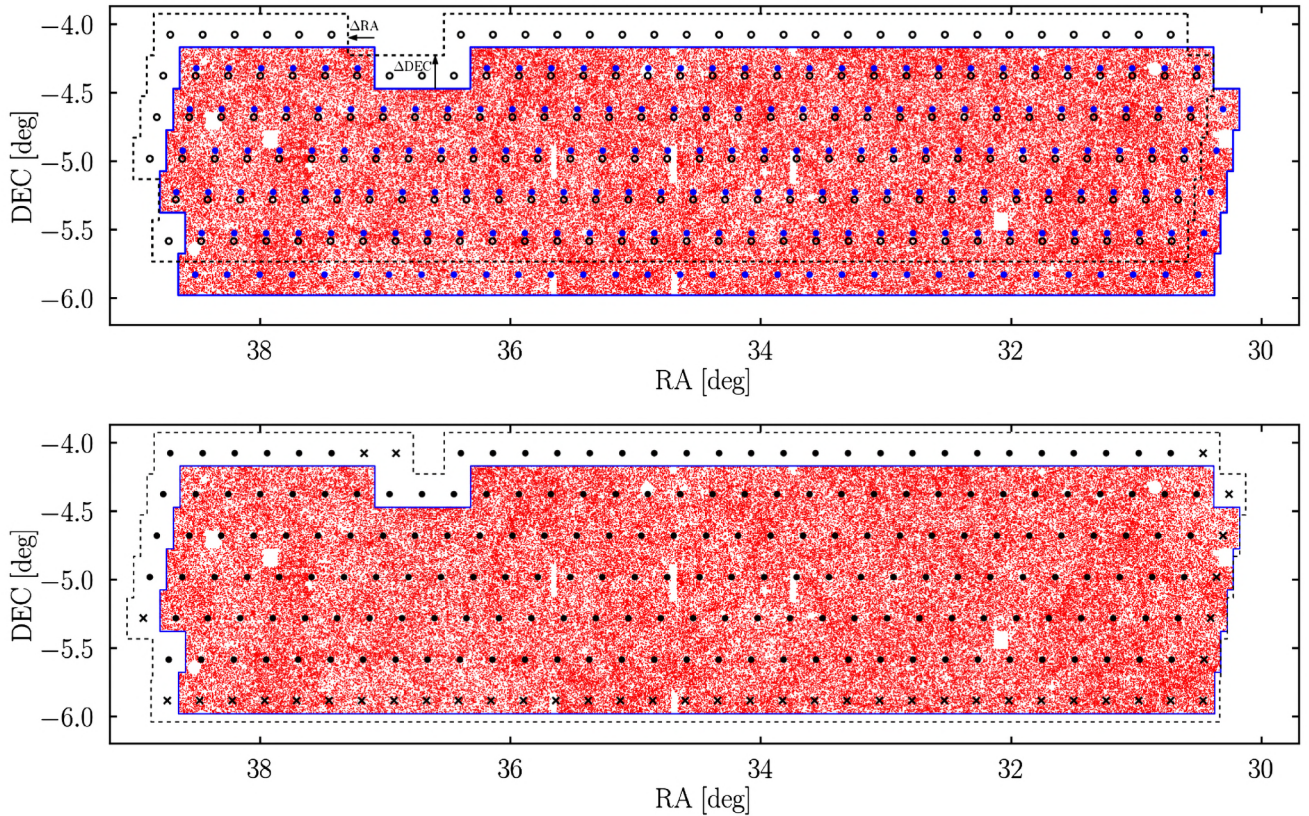
where  $w_{\text{par}}$  and  $w_{\text{targ}}$  are the 2D angular correlation functions of the parent catalogue and its sub-sample of observable pairs, respectively. We are unable to assign selection probabilities to a significant fraction of pairs at separations  $\Delta\text{RA} \lesssim 5''$  and  $\Delta\text{Dec} \lesssim 130''$  due to a combination of “slit-” and “spectra-collision” as illustrated in the left panel of Fig. 4.

Given the geometry of the problem, we were able to reduce the fraction of unobservable galaxies and galaxy-galaxy pairs by rotating the parent catalogue by  $90^\circ$ ,  $180^\circ$  and  $270^\circ$  around an axis that passes through the sample, together with random shifts of the extended spectroscopic mask. Each of the 2170 survey realisations is now characterised by a rotation angle of the corresponding parent catalogue (namely  $0^\circ$ ,  $90^\circ$ ,  $180^\circ$  and  $270^\circ$ ) and a shift of the extended spectroscopic mask in the (RA, Dec) plane. We stress here that we only rotate the parent sample while keeping the orientation of quadrants and dispersion direction of the galaxy spectra fixed; that is, the larger side of the quadrants is always aligned along the declination axis. In this way



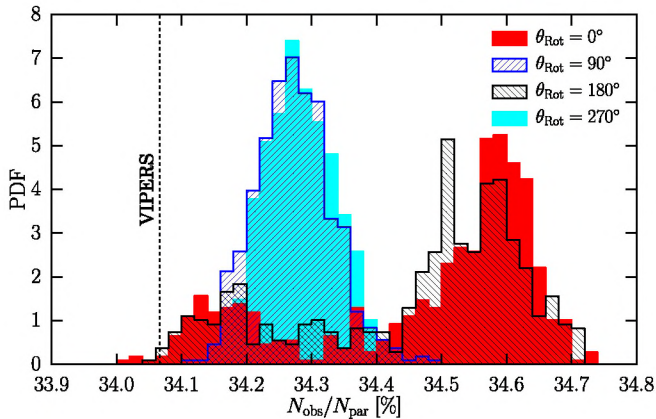


**Fig. 4.** 2D angular completeness function of galaxy-galaxy pairs (see Eq. (15)) observed in 2170 survey realisations with respect to the VIPERS parent catalogue. *Left panel:* survey realisations are obtained spatially moving the spectroscopic mask over the parent catalogue. The rectangular “shadow” at small separations is the typical footprint of VIMOS spectra. *Right panel:* as in the left panel but when also the rotations of the parent catalogue are added to make multiple survey realisations. The size of the square shadow at small pair separations in the *right panel* is the typical length of VIMOS slits and is produced by the slit collisions only.



**Fig. 5.** *Top panel:* sketch showing the border effects when multiple survey realisations are generated shifting the original VIPERS spectroscopic mask over the underlying parent catalogue (red dots). The area covered by the actual VIPERS spectroscopic mask is delimited by the blue continuous line while the corresponding VIMOS pointings are displayed as blue filled dots. A random shift of  $(\Delta RA, \Delta Dec)$  is then applied to obtain a new survey realisation. The area covered in the new realisation is shown as black dashed contour with black empty circles being the new positions of VIMOS pointings. We highlight the portion of the parent catalogues at low RA and low Dec that is not covered by the shifted mask. The effect is even more severe for the realisations obtained rotating the underlying parent catalogue. *Bottom panel:* as in the top panel but here the new survey realisation is generated shifting the “extended” spectroscopic mask (see Sect. 5). Black dots show the shifted position of the pointings in the original VIPERS spectroscopic mask while the black crosses represent the “artificial” pointings in the extended spectroscopic mask. The extended mask is large enough to fully cover also the parent catalogue rotated by  $90^\circ$ ,  $180^\circ$ , or  $270^\circ$ . In both panels a number of pointings in the shifted mask are located outside the boundaries of the parent sample. These are the pointings that only partially overlap with the parent catalogue.





**Fig. 6.** Normalised distributions of the observed fraction of VIPERS parent galaxies among 2170 targeting runs. Different colour coding and line styles differentiate runs with different rotation angles of the parent catalogue. The vertical dashed line shows the fractions of galaxies in the VIPERS-PDR2 galaxy catalogue.

we were able to assign selection probabilities to all parent galaxies and lower the fraction of unobservable galaxy-galaxy pairs to  $\sim 0.06\%$ , respectively (red dashed and blue solid lines in Fig. 3). The price to pay is that realisations with different rotation angle of the parent catalogue are not equivalent to each other in terms of the fraction of observed galaxies. In particular, rotating the parent catalogue by  $(90^\circ, 270^\circ)$  provides, on average, a number of observed galaxies that is  $\sim 1\%$  lower than the configurations with a rotation of  $(0^\circ, 180^\circ)$  as shown in Fig. 6. This is a consequence of the rectangular nature of the projected spectra and their alignment with the survey boundaries. This produces a different normalisation factor between these two sets of configurations that can be mitigated by angular up-weighting the pair counts.

Given the limited number of survey realisations we used to infer selection probabilities a small fraction of pairs remain unobserved in any realisation (we refer to them as unobservable). This introduces a systematic bias on small scales, which we do not use for RSD fitting. As discussed in Sect. 4.1, given the nature of the unobservable pairs, it is appropriate to replace  $DD^{(p)}(\theta)$  and  $DD^{(r)}(\theta)$  in Eqs. (5) and (7) with  $DD^{(\text{targ.})}(\theta)$  and  $DR^{(\text{targ.})}(\theta)$ , the number of observable galaxy-galaxy and galaxy-random pairs, respectively. In the following part, we use these quantities to compute the angular weights. Unless specified otherwise, we use the parent catalogue as a reference to estimate the systematic biases.

We treated the unobserved pointings and individual quadrants as a property of the photometric mask. Finally, we regarded the sky regions obscured by the photometric mask as a feature of the parent catalogues and imprinted the empty gaps accordingly. In particular, not imprinting the empty gaps due to unobserved pointings and quadrants in the parent catalogue would introduce a difference in the mean number of observed galaxies in different subsets of targeting runs. Indeed the gaps due to unobserved pointings in the uppermost row in the W1 field or in general those located far from the rotation axis would not be present in the configurations characterised by a rotation of the parent catalogue by  $90^\circ$  or  $270^\circ$ .

Finally, we constructed the random sample by matching the radial distribution of the VIPERS sample and imprinting the angular selection function of the parent galaxy sample, that is, applying the photometric mask. The correction scheme based on

up-weighting individual galaxies according to the local densities of parent and targeted galaxies such as the TSR weighting used in de la Torre et al. (2013) would have required including also the effect of the VIPERS spectroscopic mask. However, in our case this is not necessary, as this effect is already accounted for by using the PIP weighting. Including such a selection effect also in the random catalogue would have resulted in overweighting the pair counts.

## 6. Validation on mock catalogues

### 6.1. Consistency tests

We measured the multipole moments of the two-point correlation function from each of the 2170 survey realisations, obtained by rotating the parent catalogue and shifts of the extended spectroscopic mask, using the weighted pair counts. Each of 2170 measurements was then compared to the reference estimate obtained from the mock parent catalogue to assess the mean systematic bias and related error. These measurements are shown in the top large panels of Fig. 7 for the two redshift bins, while the bottom smaller panels show the corresponding fractional systematic bias with respect to the reference measurement. In particular, the PIP weighting scheme performs well over all scales with a systematic bias confined to the sub-percentage level on scales  $s > 1 h^{-1} \text{ Mpc}$  for all multipole moments in both redshift bins. These results are confirmed also when including the angular weights that however improve the statistical precision of the measurements.

The very small residual offset between the reference and the mean estimate among the corresponding 2170 realisations obtained using weighted pair counts is produced by the finite number of targeting runs that are used to sample the selection probabilities. A small fraction of galaxy-galaxy pairs is not observed in any of the targeting runs as shown in Fig. 3. We are therefore unable to assign selection probabilities to these objects. In particular, we can split the correlation function into two summands,

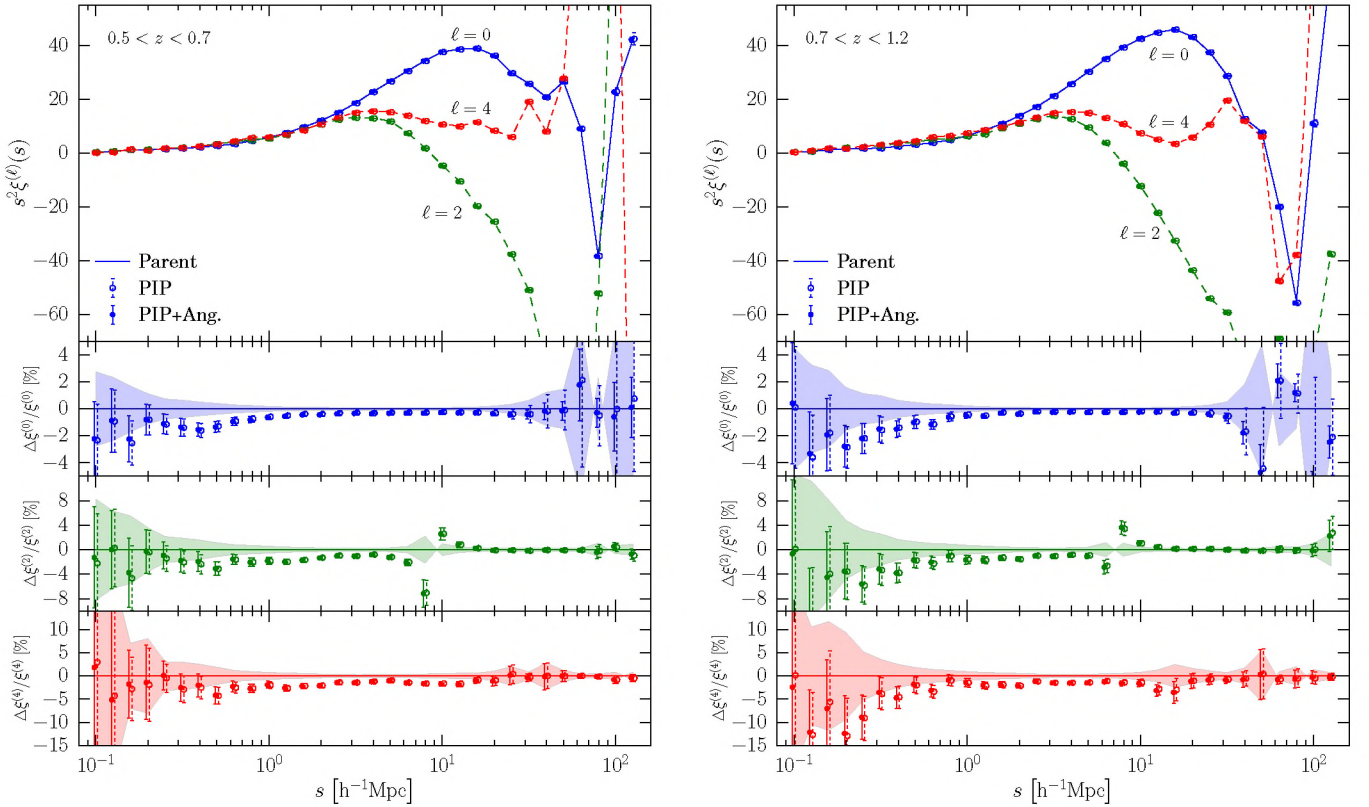
$$\xi(\mathbf{r}) = \left[ \frac{DD_{\text{obs}} - 2DR_{\text{obs}}}{RR} + 1 \right] + \left[ \frac{DD_{\text{unobs}} - 2DR_{\text{unobs}}}{RR} \right], \quad (16)$$

where the first bracket represents the contribution from the subset of observable pairs while the second one results from the unobservable pairs. We measured these quantities from a mock sample using the set of corresponding bitwise weights. It is clear from Fig. 8 that the unobservable pairs cluster in a very different way with respect to the galaxies in the full parent sample. They provide a non-negligible contribution to the overall clustering signal such that the expectation value of the estimator becomes different from that of the underlying parent sample. Indeed, the mean estimate of the two-point correlation function among 2170 survey realisation is unbiased if we limit the reference sample to only observable pairs.

### 6.2. Observational systematic bias

We quantified the observational systematic bias in the case where a set of 153 independent parent mocks is available and we have access to only one realisation of the spectroscopic observations for each parent sample, namely the one that matches the VIPERS-PDR2 observational configuration. We refer to this particular realisation as the VIPERS-like mock catalogue. We implemented the pipeline described in Sect. 5 for each of the





**Fig. 7.** *Top large panels:* measurements of the first three even multipoles of the two-point correlation function from one reference mock parent sample (lines). Points with error bars show the mean and related errors among 2170 measurements obtained using the PIP weighting scheme alone (empty markers, dashed error bars) and when supplemented with an angular up-weighting (filled markers, continuous error bars) on independent survey realisations drawn from the same mock parent sample. *Bottom small panels:* empty and filled markers display the fractional systematic bias of the corresponding measurements in the top large panels with respect to that from the reference mock parent sample. The horizontal continuous coloured lines and the shaded bands show the equivalent of the empty markers in the same panels but when the reference sample is limited to the galaxies and galaxy pairs that are targeted at least once in the 2170 survey realisations. Error bars in the *bottom panels* are obtained using the standard error propagation formula. *Left and right panels* show results from the lower- and higher-redshift bins, respectively. All measurements use data from W1 and W4 (mock) fields.

153 mock parent samples to assign selection probabilities. We measured the multipole moments from each VIPERS-like mock using the angular up-weighting and compared to the reference measurement from the corresponding parent mock to assess the observational systematic bias. The mean and related errors on such systematic biases among 153 mocks are displayed in the bottom small panels of Fig. 9 while the corresponding mean estimates and errors among 153 parent and VIPERS-like mocks are shown in the top large panels of the same figure. The measurements from the low- and high-redshift bins are shown in the left and right panels, respectively.

The new weighting scheme provides clustering measurements accurate at the sub-percentage level down to very small scales ( $\sim 0.6 h^{-1} \text{ Mpc}$ ) in both redshift ranges. The systematic bias increases at scales of  $\gtrsim 40 h^{-1} \text{ Mpc}$  remaining within  $2\sigma$  of the reference estimates. The residual systematic offset on scales of interest ( $\lesssim 50 h^{-1} \text{ Mpc}$ ) results from a combination of two effects: a) in each mock sample a small fraction of galaxy pairs remain unobserved in the ensemble of 2170 survey realisations (see Figs. 3 and 8); b) the VIPERS-like configuration is not a random realisation but rather a particular case among the 2170 survey realisations used to infer selection probabilities, namely the one characterised by a rotation angle of the parent sample of  $\theta = 0^\circ$  and no shifts (see Fig. 6). Figure 9 also shows results obtained up-weighting each galaxy by the corresponding TSR. This technique, used in previous analyses

of VIPERS-PDR2 data, performs similarly to the new method tested in this work. It is important to recall here that the TSR weighting scheme was calibrated to minimise the systematic bias on clustering estimates in mock catalogues. As such it does not assure a similar performance on real data due to possible differences between the clustering of real and simulated galaxies.

## 7. RSD fitting

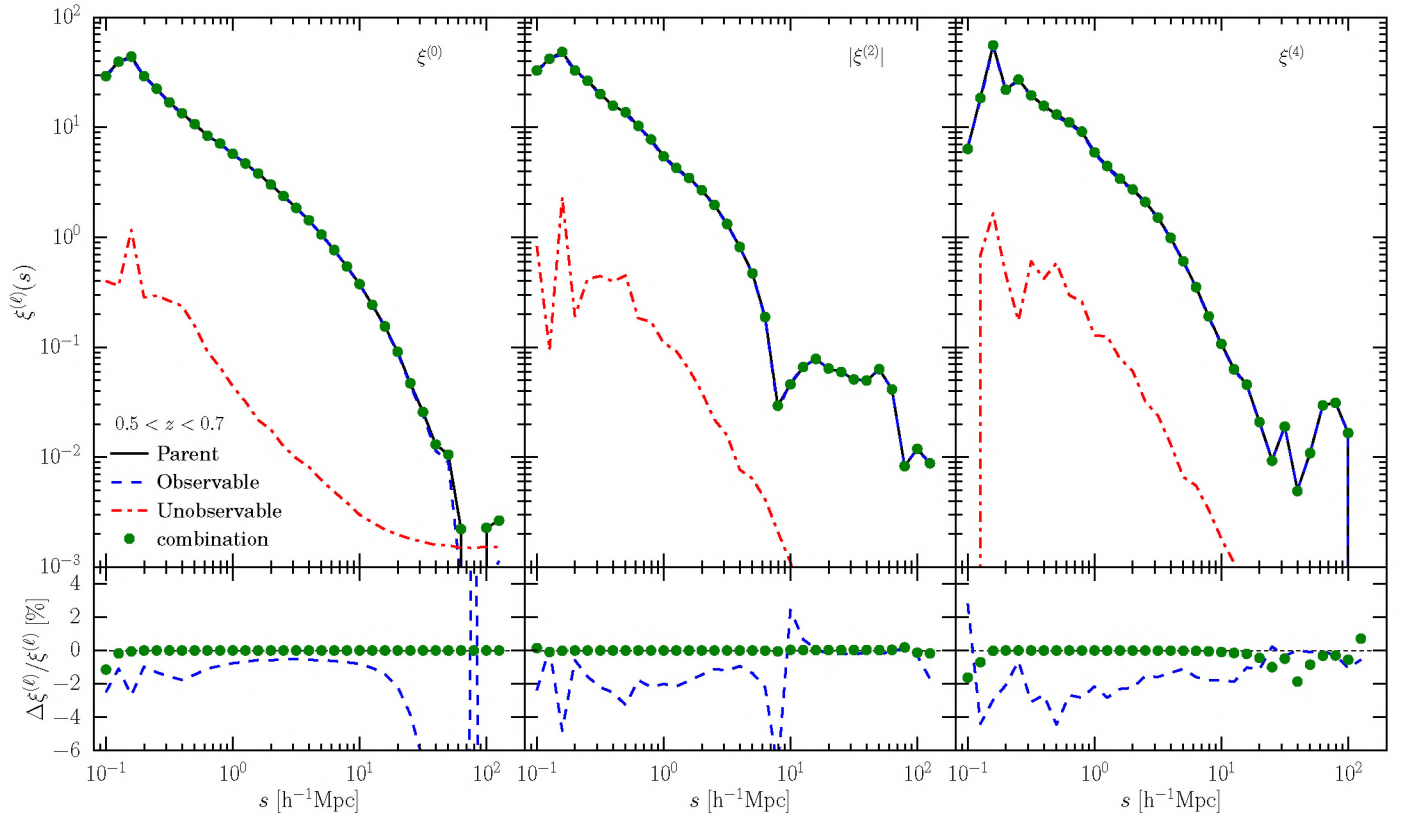
### 7.1. Theoretical modelling

We modelled the anisotropic clustering in the monopole  $\xi^{(0)}$  and quadrupole  $\xi^{(2)}$  two-point correlation functions as described in Pezzotta et al. (2017). We used the TNS model (Taruya et al. 2010) that reads in the case of biased tracers,

$$P^s(k, \mu_k) = D(k, \mu_k, \sigma_v) [b^2 P_{\delta\delta}(k) + 2\mu_k^2 f b P_{\delta\theta} + \mu_k^4 f^2 P_{\theta\theta}(k) + A(k, \mu_k, f, b) + B(k, \mu_k, f, b)], \quad (17)$$

with  $f$  and  $b$  being the growth rate and linear galaxy bias, respectively. In Eq. (17),  $P_{\delta\delta}$  is the non-linear matter power spectrum and  $P_{\delta\theta}$  and  $P_{\theta\theta}$  are density-velocity divergence and velocity divergence-velocity divergence power spectra, respectively. The correction factors  $A(k, \mu_k, f, b)$  and  $B(k, \mu_k, f, b)$  are derived using perturbation theory and provided in Taruya et al. (2010) and de la Torre & Guzzo (2012), and account for the





**Fig. 8.** *Top panels:* multipole moments measured from one mock parent catalogue (black continuous lines). The contribution to the overall clustering from the sub-samples of observable (blue dashed lines) and unobservable (red dash-dotted lines) pairs (defined respectively as those targeted at least once and the ones never targeted in the ensemble of 2170 targeting runs) as written in Eq. (16) are also shown. The combination of these two contributions is plotted as green filled markers. *Bottom panels:* fractional offset of the contribution from observable pairs and the unobservable/observable combination with respect to the reference measurement from the parent mock. This measurement refers to the low-redshift bin  $0.5 < z < 0.7$ . The measurement in the high-redshift bin  $0.7 < z < 1.2$  shows a very similar behaviour.

mode coupling between density and velocity fields. The phenomenological damping factor  $D(k\mu_k\sigma_v)$  mimics the effect of the small-scale pairwise velocity dispersion by suppressing the clustering power predicted by the “Kaiser factor” and depends on the nuisance parameter  $\sigma_v$ . We used a Lorentzian functional form for  $D(k\mu_k\sigma_v)$  as it is found to better describe the observations with respect to the theoretically predicted Gaussian damping factor (e.g. Pezzotta et al. 2017). The model in Eq. (17) is also supplemented with a second Gaussian damping factor with fixed dispersion  $\sigma_z$  to account for the effect of VIPERS redshift errors on clustering measurements.

The model in Eq. (17) depends on four fitting parameters ( $f, b, \sigma_8, \sigma_v$ ). However, we provide measurements of the derived parameters  $f\sigma_8$  and  $b\sigma_8$  as  $\sigma_8$ , the normalisation of the linear matter power spectrum  $P_{\delta\delta}^{\text{lin}}$ , is degenerate with the growth rate parameter  $f$  and the linear bias factor  $b$ .

The linear matter power spectrum  $P_{\delta\delta}^{\text{lin}}$  is obtained using the Code for Anisotropies in the Microwave Background (Lewis et al. 2000, CAMB) that is combined with HALOFIT (Smith et al. 2003; Takahashi et al. 2012) to predict the non-linear matter power spectrum  $P_{\delta\delta}$ . The density-velocity divergence  $P_{\delta\theta}$  and velocity divergence-velocity divergence  $P_{\theta\theta}$  power spectra cannot be measured from data directly. They can be predicted by either using perturbation theory or by means of empirical fitting functions calibrated on numerical simulations (e.g. Jennings et al. 2011). Perturbation theory however breaks down at scales accessible in VIPERS. We therefore used the improved fitting functions described in Bel et al. (in prep.),

$$P_{\delta\theta}(k) = \left[ P_{\delta\delta}^{\text{lin}}(k) P_{\delta\delta}(k) \exp\left(-\frac{k}{k_{\delta\theta}^{\text{cut}}}\right) \right]^{1/2}, \quad (18a)$$

$$P_{\theta\theta}(k) = \left[ P_{\delta\delta}^{\text{lin}}(k) \exp\left(-\frac{k}{k_{\theta\theta}^{\text{cut}}}\right) \right]. \quad (18b)$$

In Eq. (18),  $k_{\delta\theta}^{\text{cut}}$  and  $k_{\theta\theta}^{\text{cut}}$  are defined as

$$k_{\delta\theta}^{\text{cut}} = \frac{1}{2.972} \sigma_8^{-2.034}, \quad (19a)$$

$$k_{\theta\theta}^{\text{cut}} = \frac{1}{1.906} \sigma_8^{-2.163}, \quad (19b)$$

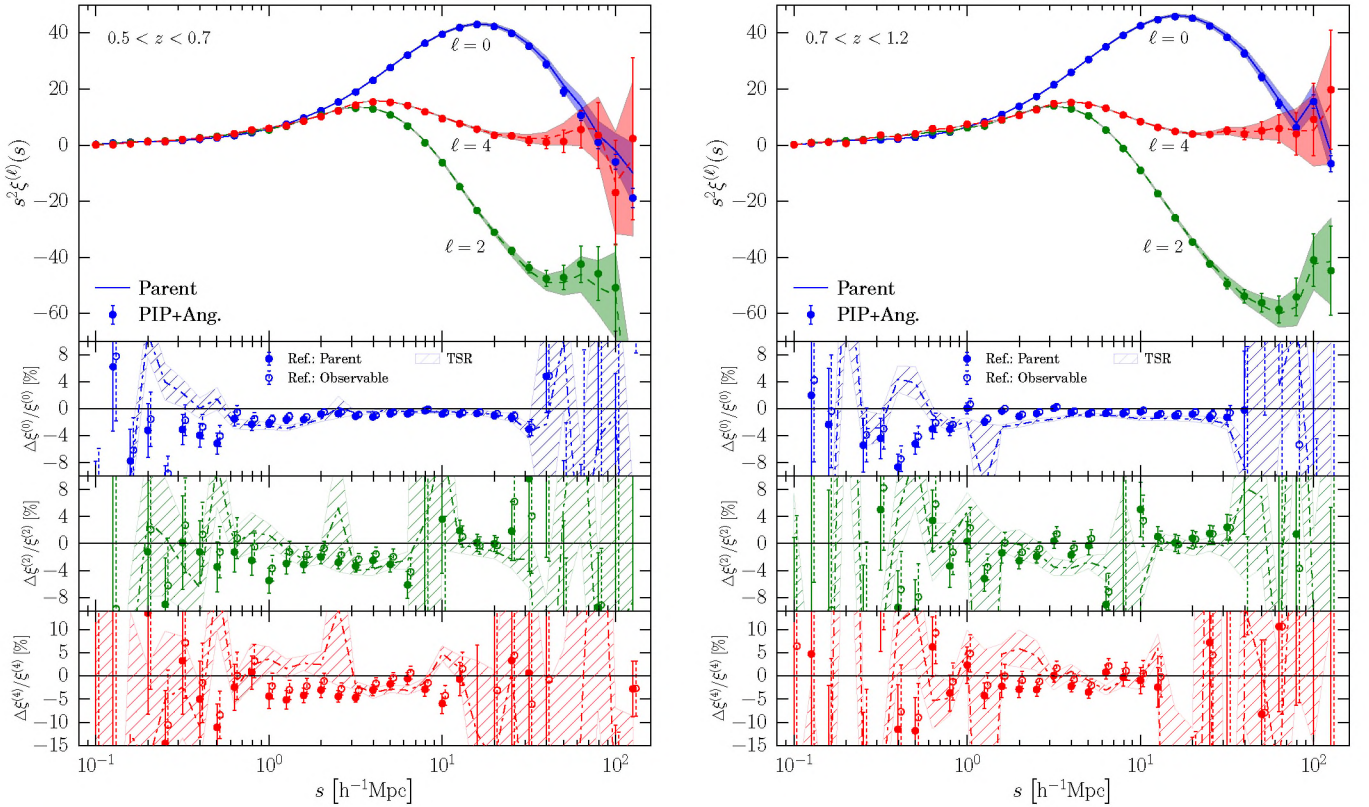
with  $\sigma_8$  being the amplitude of the linear matter power spectrum. We note that in our model,  $\sigma_8$  controls the level of non-linearity (within HALOFIT) in the matter non-linear density-velocity divergence and velocity divergence-velocity divergence power spectra that enter the RSD model of Eq. (17).

## 7.2. Fitting method and data covariance matrix

The measured monopole and quadrupole are simultaneously fitted with the TNS model to estimate the fitting parameters using the Monte-Carlo Markov chain (MCMC) technique. The MCMC algorithm explores the posterior distribution in the parameter space constrained by the data likelihood and parameter priors. The data likelihood is,

$$-2 \ln \mathcal{L} = \chi^2(\theta_p) = \sum_{i,j} \Delta_i(\theta_p) C_{ij}^{-1} \Delta_j(\theta_p), \quad (20)$$





**Fig. 9.** *Top panels:* mean estimates and related errors of multipole moments of the two-point correlation function from the set of 153 mock parent samples (lines with shaded bands) and the corresponding VIPERS-like mocks obtained using the PIP and angular up-weighting method (points with error-bars). *Bottom panels:* mean fractional systematic bias of measurements from VIPERS-like mocks with respect to the ones from the underlying parent samples (filled points with solid error-bars). In the *bottom small panels* we also display the case when the sub-sample of observable pairs is used as reference (empty markers with dashed error-bars). Measurements obtained using the TSR weighting scheme are plotted for comparison (dash-dotted lines with hatched areas). Error bars in the bottom panels quantify the scatter of the systematic offsets among 153 mocks. *Left and right panels* show the measurements in the low  $0.5 < z < 0.7$  and high  $0.7 < z < 1.2$  redshift bins, respectively.

where  $\theta_p$  denotes the set of fitting parameters,  $\Delta_i$  is the discrepancy between the data and model prediction in bin  $i$  and  $C_{ij}^{-1}$  is the precision matrix, that is, the inverse of the data covariance matrix  $C_{ij}$ . We fit the monopole  $s^2 \xi^{(0)}$  and quadrupole  $s^2 \xi^{(2)}$  of the two-point correlation functions simultaneously and accounted for their cross-covariance in the data covariance matrix.

The covariance matrices  $C_{ij}$  were estimated using the set of 153 VIPERS-like mocks. Noise in the covariance matrix is amplified when inferring the precision matrix using  $C_{ij}$  and leads to a biased estimate of the precision matrix. We corrected for this bias by means of the corrective factor provided in Percival et al. (2014). The correlation matrices, that is,  $R_{ij} = C_{ij}/(C_{ii}C_{jj})^{1/2}$ , for the two redshift bins and restricted to the range of fitting scales used here are shown in Fig. 10.

The robustness of the data analysis method has already been tested in Pezzotta et al. (2017). We therefore focus on repeating the analysis using only the range of fitting scales adopted in Pezzotta et al. (2017) to obtain the reference estimates of the  $f\sigma_8$  parameter, that is, minimum and maximum scales fixed at  $s_{\min} = 5 \text{ h}^{-1} \text{ Mpc}$  and  $s_{\max} = 50 \text{ h}^{-1} \text{ Mpc}$ , respectively. In particular, we fit the mean estimates of  $s^2 \xi^{(\ell)}$  for the monopole  $\ell = 0$  and quadrupole  $\ell = 2$  from the mock catalogues with the TNS model and obtained a systematic offset, with respect to the fiducial values, of

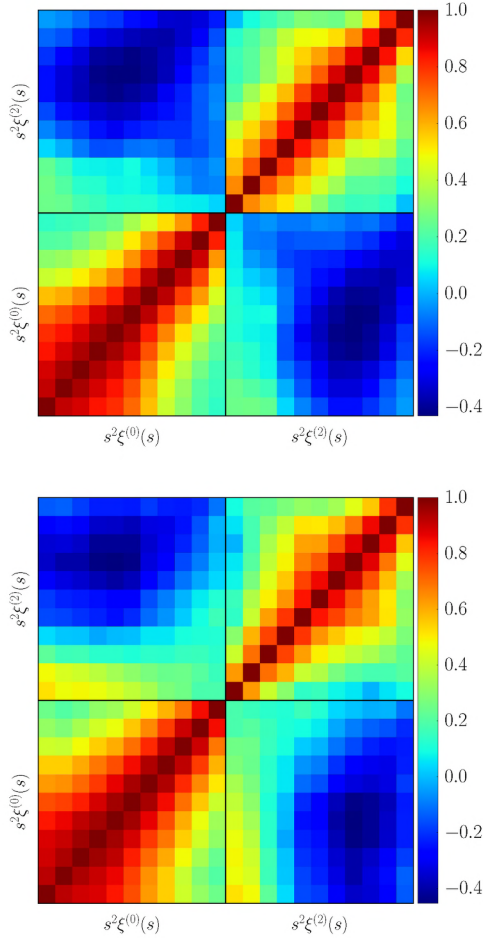
$$\begin{aligned} \Delta(f\sigma_8)(z = 0.60) &= 0.009 \pm 0.015 \\ \Delta(f\sigma_8)(z = 0.86) &= -0.006 \pm 0.012. \end{aligned}$$

These estimates are unbiased compared to the expected values of  $f\sigma_8(z)$  in the mock fiducial cosmology. Moreover our measurements are also compatible with estimates obtained in Pezzotta et al. (2017),  $\Delta(f\sigma_8)(z = 0.60) = 0.019 \pm 0.012$  and  $\Delta(f\sigma_8)(z = 0.86) = -0.018 \pm 0.011$ , within  $1\sigma$ . The marginalised one- and two-dimensional posterior likelihoods are shown in Fig. 11. For comparison we also show, in the same figure, the results obtained by Pezzotta et al. (2017) using the same set of mocks.

## 8. Growth rate measurements

To correct the measurements of the two-point correlation function from the VIPERS-PDR2 galaxy catalogue we followed the same procedure adopted on mock catalogues, including calculating the PIP weights using both rotations to the parent catalogue and shifts of the extended spectroscopic mask. As VIPERS parent catalogue we used the photometric catalogue from the CFHTLS W1 and W4 fields, from which VIPERS targets were drawn, restricted to the area covered by the VIPERS observations. However, unlike mock samples the VIPERS parent catalogue contains  $N_c = 449$  compulsory targets that do not enter the maximisation of the number of slits. Although a negligible fraction, we accounted for these objects when generating multiple survey realisations unless they fall inside the empty gaps between VIMOS quadrants. As anticipated in Sect. 2, we used only galaxies with quality flags 2–9 (inclusive) corresponding to





**Fig. 10.** Data correlation matrices  $R_{ij} = C_{ij}/\sqrt{C_{ii}C_{jj}}$  estimated using the set of 153 VIPERS-like mock catalogues in the low-  $0.5 < z < 0.7$  (top panel) and high-redshift bins  $0.7 < z < 1.2$  (bottom panel).

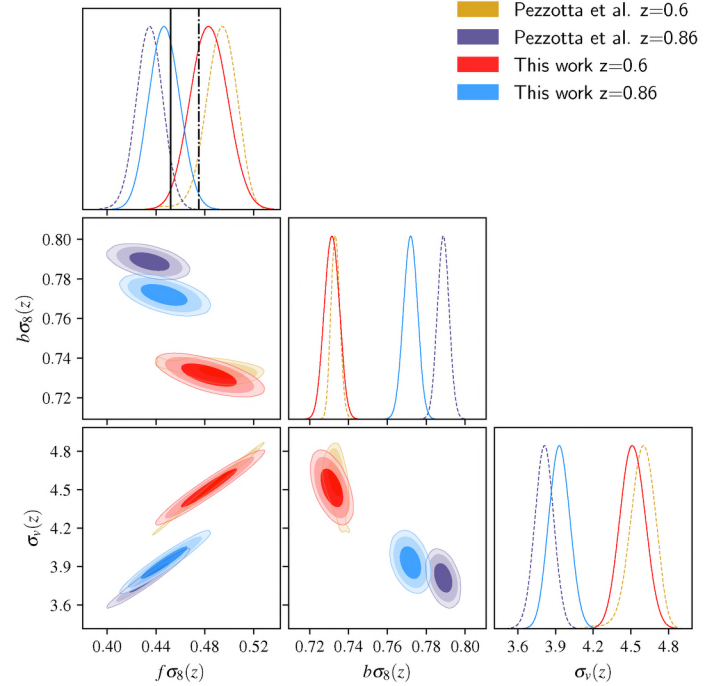
a sample with a redshift confirmation rate of 96.1%. The effect of redshift failures is not accounted for when computing the PIP weights. We therefore corrected for the effect of redshift failures by up-weighting each galaxy in the VIPERS-PDR2 catalogue by the corresponding SSR as described in Sect. 4.2.

We fit the monopole  $s^2\xi^{(0)}$  and quadrupole  $s^2\xi^{(2)}$  of the two-point correlation function between  $s_{\min} = 5 h^{-1}$  Mpc and  $s_{\max} = 50 h^{-1}$  Mpc with the TNS model in Eq. (17) supplemented with a second Gaussian damping factor with width fixed to the VIPERS spectroscopic redshift error  $\sigma_z/(1+z) = 0.00054$ . The measured values for the derived parameter  $f\sigma_8$  are,

$$f\sigma_8(z = 0.60) = 0.49 \pm 0.12$$

$$f\sigma_8(z = 0.86) = 0.46 \pm 0.09.$$

These values are compatible within  $1-\sigma$  with estimates from Pezzotta et al. (2017), namely  $f\sigma_8(z = 0.60) = 0.55 \pm 0.12$  and  $f\sigma_8(z = 0.86) = 0.40 \pm 0.11$ , who used the same datasets and theoretical prescriptions for RSD modelling. Furthermore our measurements are also consistent within the error bars with the ones obtained with alternative methods such as a combination of RSD and galaxy-galaxy lensing in de la Torre et al. (2017) or the one using a sample of luminous blue galaxies in VIPERS as done in Mohammad et al. (2018). The best-fit models corresponding to the results in Fig. 12 are displayed in Fig. 13 along with the measurements of the monopole  $s^2\xi^{(0)}$  and quadrupole  $s^2\xi^{(2)}$  moments of the two-point correlation function using the VIPERS-PDR2



**Fig. 11.** One- and two-dimensional marginalised posterior likelihoods of the derived parameter  $f\sigma_8$ ,  $b\sigma_8$  and the nuisance parameter  $\sigma_v$  resulting from the analysis of the mean clustering estimates obtained from 153 VIPERS-like mock catalogues using the method in Sect. 4.1. Fits are performed with TNS model between a minimum fitting scale of  $s_{\min} = 5 h^{-1}$  Mpc up to a maximum scale of  $s_{\max} = 50 h^{-1}$  Mpc. For comparison we have also over-plotted results obtained in Pezzotta et al. (2017) using the same set of mock samples and fitting method. Vertical dash-dotted and solid lines correspond to the expected values of  $f\sigma_8$  at  $z = 0.6$  and  $z = 0.86$ , respectively.

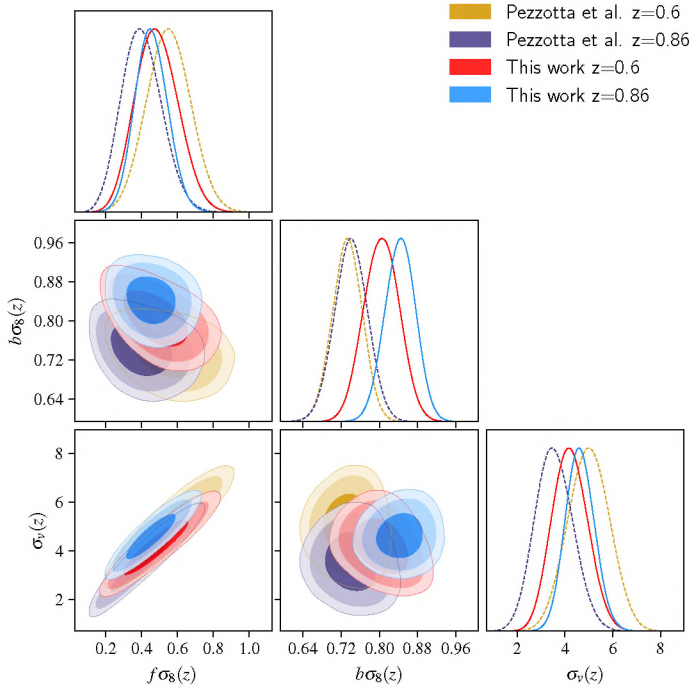
galaxy sample (points with error-bars) and VIPERS-like mocks (cyan lines).

## 9. Summary and conclusions

We corrected the clustering estimates from the VIPERS-PDR2 galaxy sample using the PIP method described in Bianchi & Percival (2017). This technique was supplemented with the angular up-weighting scheme proposed in Percival & Bianchi (2017) to improve the statistical precision of the measurements. The PIP method relies on up-weighting the pair-counts based on the corresponding selection probabilities. These probabilities were inferred empirically by generating multiple survey realisations from a parent catalogue and counting the number of times a given pair is observed. To compare the performance of this new technique with the results obtained in Pezzotta et al. (2017) we split the redshift range probed by VIPERS into two bins spanning  $0.5 < z < 0.7$  and  $0.7 < z < 1.2$ . The following considerations equally apply to both redshift bins.

Given the features of the VIPERS targeting algorithm and the limited extension of the VIPERS parent mocks, we generated multiple (2170) VIPERS realisations from each parent sample by spatially moving the spectroscopic mask. To assign selection probabilities to galaxy pairs with a reasonable amount of computation time we also rotated the parent catalogue in each targeting run. The price to pay is that survey realisations with different rotations of the parent sample are not fully equivalent to each other producing a “normalisation problem” for the weighted pair



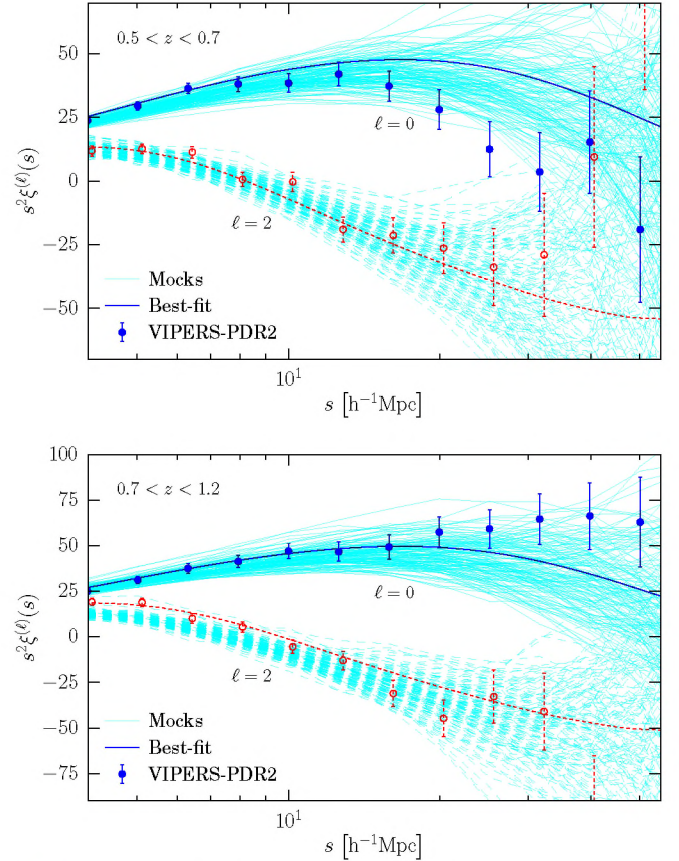


**Fig. 12.** As in Fig. 11 but now fitting the monopole  $\xi^{(0)}$  and quadrupole  $\xi^{(2)}$  measured from the VIPERS-PDR2 spectroscopic sample. Again, results of RSD fitting in Pezzotta et al. (2017) are also plotted.

counts. We mitigated for this problem supplementing the PIP technique with the angular up-weighting method. A negligible mean systematic bias was found comparing clustering measurements from each of 2170 survey realisations with the reference measurement from the parent catalogue. Nevertheless, we have shown that this bias is produced by the very small fraction of galaxy pairs unobserved in  $N_{\text{runs}} = 2170$  survey realisations. Indeed these pairs are not randomly distributed but rather exhibit a small-scale clustering.

To assess the observational systematic bias on clustering measurements, we selected, for each parent mock, only the survey realisation obtained with actual VIPERS observational setup, that is no rotation of the parent sample and no shift in the spectroscopic mask. We found a mean fractional systematic bias among 153 mock samples to be below the percentage level. We argue that such a small offset results from a combination of two effects: a) we are unable to assign selection probabilities to a small fraction of pairs that cannot be observed using only 2170 survey realisations, referred to as unobservable pairs; and b) the VIPERS-like mock is a particular configuration among the 2170 realisations used to infer the selection probabilities. Our tests using mocks catalogues have shown the new method to be a valid and robust way to correct for missing targets in VIMOS observations.

We tested the impact of these corrections on estimates of the growth rate of structure times the amplitude of dark matter density fluctuations  $f\sigma_8$ . In particular we fitted the mean estimates of the corrected monopole and quadrupole among 153 VIPERS-like mocks with the TNS model on scales  $5 \text{ h}^{-1} \text{ Mpc} < s < 50 \text{ h}^{-1} \text{ Mpc}$ . The analysis provided un-biased estimates of the fitting parameter  $f\sigma_8$  that are fully consistent with those obtained in Pezzotta et al. (2017) using the same configuration of fitting scales and theoretical model. The measurements made using the new technique are slightly closer to the expected values, but the difference is within the expected errors. This provides further confirmation of the robustness of previous RSD



**Fig. 13.** Monopole  $s^2\xi^{(0)}$  and quadrupole  $s^2\xi^{(2)}$  moments of the two-point correlation function measured from VIPERS-PDR2 galaxy sample using the weighted pair counts as described in Sect. 4.1 (points with error-bars). Diagonal errors are estimated using the set of 153 VIPERS-like mocks. Cyan lines show the measurements from individual VIPERS-like mocks. The best-fit models corresponding to the results in Fig. 12 are also displayed as solid blue and dashed red lines. *Top and bottom panels* show results from the low- and high-redshift bins, respectively.

analyses in VIPERS. However we stress here the fact that while the correction scheme adopted in previous VIPERS works (e.g. de la Torre et al. 2017; Pezzotta et al. 2017; Mohammad et al. 2018) relied on a fine-tuned parametric approach calibrated on mock catalogues to minimise the observational systematic bias, the new technique proposed in Bianchi & Percival (2017) and Percival & Bianchi (2017) is exact and is self-contained, using only the data itself.

Finally, we applied this method to correct the measurements of the two-point correlation function using the VIPERS-PDR2 galaxy catalogue. When dealing with data we have accounted for the effect of redshift failures by means of the so called “spectroscopic success rate” (SSR). We also took into account the presence of a small fraction of compulsory targets in the parent sample. Both these features were not reproduced in the mock samples. The measured monopole and quadrupole moments of the two-point correlation functions were fitted with the TNS model to estimate the  $f\sigma_8$  parameter at the effective redshifts of the two redshift bins. Our measurements are in agreement within  $1\text{-}\sigma$  with previous measurements by Pezzotta et al. (2017), de la Torre et al. (2017) and Mohammad et al. (2018) at the same redshifts.

In future work, we will improve upon this analysis using the method of Percival & Bianchi (2017) to include angular



clustering measurements from the full CFHTLS sample. By using a combination of the angular and 3D clustering measurements, we hope to observe baryon acoustic oscillations, as well as to improve on the current RSD measurements.

**Acknowledgements.** We thank Michael Wilson for useful comments on this work. We acknowledge the crucial contribution of the ESO staff for the management of service observations through which the VIPERS survey was built. In particular, we are deeply grateful to M. Hilker for his constant help and support of this programme. Italian participation to VIPERS has been funded by INAF through PRIN 2008, 2010, 2014 and 2015 programs. LG, FGM, BRG and JB acknowledge support from the European Research Council through grant n. 291521. DB and WJP acknowledge support from the European Research Council through grant n. 614030. OLF acknowledges support from the European Research Council through grant n. 268107. SDLT acknowledges the support of the OCEVU Labex (ANR-11-LABX-0060) and the A\*MIDEX project (ANR-11-IDEX-0001-02) funded by the “Investissements d’Avenir” French government programme managed by the ANR. RT acknowledges financial support from the European Research Council through grant n. 202686. AP, KM, and JK have been supported by the National Science Centre (grants UMO-2012/07/B/ST9/04425 and UMO-2013/09/D/ST9/04030). EB, FM and LM acknowledge the support from grants ASI-INAf I/023/12/0 and PRIN MIUR 2010-2011. TM and SA acknowledge financial support from the ANR Spin(e) through the French grant ANR-13-BS05-0005. The Big MultiDark Database used in this paper and the web application providing online access to it were constructed as part of the activities of the German Astrophysical Virtual Observatory as result of a collaboration between the Leibniz-Institute for Astrophysics Potsdam (AIP) and the Spanish MultiDark Consolider Project CSD2009-00064. The Bolshoi and MultiDark simulations were run on the NASA’s Pleiades super-computer at the NASA Ames Research Center.

## References

- Bianchi, D., & Percival, W. J. 2017, *MNRAS*, **472**, 1106
- Bianchi, D., Burden, A., & Percival, W. J. 2018, *MNRAS*, **481**, 2338B
- Bottini, D., Garilli, B., Maccagni, D., et al. 2005, *PASP*, **117**, 996
- Dawson, K. S., Kneib, J.-P., Percival, W. J., et al. 2016, *AJ*, **151**, 44
- de la Torre, S., & Guzzo, L. 2012, *MNRAS*, **427**, 327
- de la Torre, S., & Peacock, J. A. 2013, *MNRAS*, **435**, 743
- de la Torre, S., Guzzo, L., Peacock, J. A., et al. 2013, *A&A*, **557**, A54
- de la Torre, S., Jullo, E., Giocoli, C., et al. 2017, *A&A*, **608**, A44
- Garilli, B., Guzzo, L., Scodreggio, M., et al. 2014, *A&A*, **562**, A23
- Guzzo, L., Scodreggio, M., Garilli, B., et al. 2014, *A&A*, **566**, A108
- Jennings, E., Baugh, C. M., & Pascoli, S. 2011, *MNRAS*, **410**, 2081
- Kaiser, N. 1987, *MNRAS*, **227**, 1
- Klypin, A., Yepes, G., Gottlöber, S., Prada, F., & Heß, S. 2016, *MNRAS*, **457**, 4340
- Landy, S. D., & Szalay, A. S. 1993, *ApJ*, **412**, 64
- Le Fèvre, O., Saisse, M., & Mancini, D., et al. 2003, *Proc. SPIE*, **4841**, 1670
- Lewis, A., Challinor, A., & Lasenby, A. 2000, *ApJ*, **538**, 473
- Mohammad, F. G., Granett, B. R., Guzzo, L., et al. 2018, *A&A*, **610**, A59
- Percival, W. J., & Bianchi, D. 2017, *MNRAS*, **472**, L40
- Percival, W. J., Ross, A. J., Sánchez, A. G., et al. 2014, *MNRAS*, **439**, 2531
- Pezzotta, A., de la Torre, S., Bel, J., et al. 2017, *A&A*, **604**, A33
- Scodreggio, M., Franzetti, P., Garilli, B., Le Fèvre, O., & Guzzo, L. 2009, *The Messenger*, **135**, 13
- Scodreggio, M., Guzzo, L., & Garilli, B. 2018, *A&A*, **609**, A84
- Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, *MNRAS*, **341**, 1311
- Takahashi, R., Sato, M., Nishimichi, T., Taruya, A., & Oguri, M. 2012, *ApJ*, **761**, 152
- Taruya, A., Nishimichi, T., & Saito, S. 2010, *Phys. Rev. D*, **82**, 063522
- <sup>1</sup> INAF – Osservatorio Astronomico di Brera, Via Brera 28, 20122 Milano – via E. Bianchi 46, 23807 Merate, Italy  
e-mail: faizan.mohammad@brera.inaf.it
- <sup>2</sup> Università degli Studi di Milano, via G. Celoria 16, 20133 Milano, Italy
- <sup>3</sup> Institute for Cosmology and Gravitation, University of Portsmouth, 1-8 Burnaby Rd, Portsmouth PO1 3, UK
- <sup>4</sup> Department of Physics and Astronomy, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada
- <sup>5</sup> Perimeter Institute for Theoretical Physics, 31 Caroline St. North, Waterloo, ON N2L 2Y5, Canada
- <sup>6</sup> Aix-Marseille Université, Université Toulon, CNRS, CPT, Marseille, France
- <sup>7</sup> Dipartimento di Matematica e Fisica, Università degli Studi Roma Tre, via della Vasca Navale 84, 00146 Roma, Italy
- <sup>8</sup> INFN, Sezione di Roma Tre, via della Vasca Navale 84, 00146 Roma, Italy
- <sup>9</sup> INAF – Osservatorio Astronomico di Roma, via Frascati 33, 00040 Monte Porzio Catone (RM), Italy
- <sup>10</sup> Aix-Marseille Université, CNRS, CNES, LAM, Marseille, France
- <sup>11</sup> Dipartimento di Fisica e Astronomia – Alma Mater Studiorum Università di Bologna, via Gobetti 93/2, 40129 Bologna, Italy
- <sup>12</sup> INFN, Sezione di Bologna, viale Berti Pichat 6/2, 40127 Bologna, Italy
- <sup>13</sup> INAF – Osservatorio Astronomico di Bologna, via Gobetti 93/3, 40129 Bologna, Italy
- <sup>14</sup> INAF – Istituto di Astrofisica Spaziale e Fisica Cosmica Milano, via Bassini 15, 20133 Milano, Italy
- <sup>15</sup> INAF – Osservatorio Astrofisico di Torino, 10025 Pino Torinese, Italy
- <sup>16</sup> Laboratoire Lagrange, UMR7293, Université de Nice Sophia Antipolis, CNRS, Observatoire de la Côte d’Azur, 06300 Nice, France
- <sup>17</sup> Institute of Physics, Jan Kochanowski University, ul. Świetokrzyska 15, 25-406 Kielce, Poland
- <sup>18</sup> National Centre for Nuclear Research, ul. Hoza 69, 00-681 Warszawa, Poland
- <sup>19</sup> Aix-Marseille Université, Jardin du Pharo, 58 bd Charles Livon, 13284 Marseille Cedex 7, France
- <sup>20</sup> IRAP, 9 av. du colonel Roche, BP 44346, 31028 Toulouse Cedex 4, France
- <sup>21</sup> Astronomical Observatory of the Jagiellonian University, Orla 171, 30-001 Cracow, Poland
- <sup>22</sup> School of Physics and Astronomy, University of St Andrews, St Andrews KY16 9SS, UK
- <sup>23</sup> INAF – Istituto di Astrofisica Spaziale e Fisica Cosmica Bologna, via Gobetti 101, 40129 Bologna, Italy
- <sup>24</sup> INAF – Istituto di Radioastronomia, via Gobetti 101, 40129 Bologna, Italy
- <sup>25</sup> Canada-France-Hawaii Telescope, 65–1238 Mamalahoa Highway, Kamuela, HI 96743, USA
- <sup>26</sup> Department of Astronomy, University of Geneva, ch. d’Ecogia 16, 1290 Versoix, Switzerland
- <sup>27</sup> INAF – Osservatorio Astronomico di Trieste, via G. B. Tiepolo 11, 34143 Trieste, Italy
- <sup>28</sup> Department of Astronomy & Physics, Saint Mary’s University, 923 Robie Street, Halifax Nova Scotia B3H 3C3, Canada